IT & C

ISSN 2821 - 8469, ISSN – L 2821 - 8469, Volumul 1, Numărul 2, Decembrie 2022

Securitatea informațiilor în lucrul cu megadate pe Internet

Nicolae Sfetcu

Pentru a cita acest articol: Sfetcu, Nicolae (2022), Securitatea informațiilor în lucrul cu megadate pe Internet, *IT & C*, **1**:2, 74-84, DOI: 10.58679/IT75075, https://www.internetmobile.ro/securitatea-informatiilor-in-lucrul-cu-megadate-pe-internet/

Publicat online: 16.09.2022

Securitatea informațiilor în lucrul cu megadate pe Internet

Nicolae Sfetcu

Rezumat

Securitatea megadatelor (big data) presupune aderarea la conceptele de comportament etic corect și greșit în ceea ce privește datele, în special datele cu caracter personal. Etica big data pune accentul pe colectorii și diseminatorii de date structurate sau nestructurate. Securitatea și confidențialitatea informațiilor este susținută, la nivelul UE, de o amplă documentație, prin care se încearcă să se găsească soluții concrete pentru maximizarea valorii informațiilor fără a sacrifica drepturile fundamentale ale omului. Autoritatea Europeană pentru Protecția Datelor (AEPD) sprijină dreptul la viață privată și dreptul la protecția datelor cu caracter personal în respectul demnității umane.

Cuvinte cheie: securitatea informațiilor, securitate, informații, megadate, big data, Internet, confidențialitate, GDPR

Abstract

Big data security involves adhering to the concepts of right and wrong ethical behavior with respect to data, especially personal data. Big data ethics focuses on the collectors and disseminators of structured or unstructured data. Information security and privacy is supported, at EU level, by extensive documentation, which seeks to find concrete solutions to maximize the value of information without sacrificing fundamental human rights. The European Data Protection Authority (EDPS) supports the right to privacy and the right to the protection of personal data respecting human dignity.

Keywords: information security, security, information, megadata, big data, Internet, privacy, GDPR

IT & C, Volumul 1, Numărul 2, Decembrie 2022, pp. 74-84
ISSN 2821 - 8469, ISSN – L 2821 - 8469
URL: https://www.internetmobile.ro/securitatea-informatiilor-in-lucrul-cu-megadate-pe-internet/

Securitatea informațiilor presupune aderarea la conceptele de comportament corect și greșit în ceea ce privește datele, în special datele cu caracter personal. Etica megadatelor pune accentul pe colectorii și diseminatorii de date structurate sau nestructurate.

Securitatea și confidențialitatea informațiilor este susținută, la nivelul UE, de o amplă documentație, prin care se încearcă să se găsească soluții concrete pentru maximizarea valorii Big Data fără a sacrifica drepturile fundamentale ale omului. Autoritatea Europeană pentru Protecția Datelor (AEPD) sprijină dreptul la viață privată și dreptul la protecția datelor cu caracter personal în respectul demnității umane. Conform acestor documente, trebuie să se depășească conflictul conceptual dintre viața privată și Big Data,și între intimitate și inovație. Este esențial să se identifice modalitățile de includere a dimensiunii etice în conceperea inovațiilor. (European Economic and Social Committee 2017)

Conform noului Regulament al UE 2016/679, operatorii de date trebuie să pună în aplicare măsurile de confidențialitate și tehnologiile de îmbunătățire a confidențialității în momentul determinării modalităților de procesare și al procesării în sine. Prin ENISA75 au fost identificate multe strategii de confidențialitate prin design (minimizarea datelor, ascunderea datelor cu caracter personal și interconexiunile acestora, prelucrarea separată a datelor cu caracter personal, alegerea celui mai înalt nivel de agregare, transparența, monitorizarea, politica de confidențialitate, aspecte legale).

O modalitate de bază pentru coexistența pașnică dintre exploatarea Big Data și protecția datelor este *controlul*, de către utilizator, a datelor personale, ceea ce conduce la transparență și încredere între utilizatori și furnizorii de servicii digitale. După cum s-a subliniat în evaluarea impactului GDPR,

> "Construirea încrederii în mediul online este esențială pentru dezvoltarea economică. Lipsa de încredere îi face pe consumatori să ezite să cumpere online și să adopte noi servicii, inclusiv serviciile publice de e-guvernare. Dacă nu este abordată, această lipsă de încredere va continua să încetinească dezvoltarea utilizărilor inovatoare ale noilor tehnologii, să acționeze ca un obstacol în calea creșterii economice și să blocheze sectorul public în a profita de beneficiile potențiale ale digitalizării serviciilor sale."

În cazul Big Data modelele tradiționale de *consimțământ* sunt insuficiente și depășite. ”Consimțământul ar trebui să fie suficient de granular pentru a acoperi toate procesele diferite și scopurile de prelucrare și reutilizare a datelor cu caracter personal.” (European Economic and Social Committee 2017)

O problemă specială este *portabilitatea* datelor, susținută la nivelul UE de AEPD în Avizul 7/2015, (MORO 2016) unde se impune garantarea dreptului cetățenilor de a accesa și corecta datele personale printr-un control extins. Portabilitatea datelor poate ajuta la creșterea gradului de conștientizare și control al consumatorilor prin transferul între servicii online.

AEPD consideră că datele cu caracter personal ar trebui să fie tratate la fel ca alte resurse importante, precum petrolul, unde tranzacționarea are loc între părți la fel de bine informate (simetria informațională). În realitate, piața informațiilor cu caracter personal are un caracter de asimetrie informațională, nefiind nici transparentă, nici echitabilă, clienții nefiind compensați pentru informațiile personale pe care le oferă. Astfel, portabilitatea datelor ar încuraja un mediu mai competitiv între beneficiarii acestor date, utilizatorii având posibilitatea să aleagă cui oferă atele personale.

O altă abordare pune în discuție *stocarea* datelor cu caracter personal, cu posibilitatea pentru utilizator de a acorda sau retrage consimțământul pentru datele sale personale. (MORO 2016) (DG Connect 2015) Stocarea datelor cu caracter personal implică un ”concept cadru și o implementare arhitecturală care transferă achiziția și controlul datelor de la un model de date distribuit la un model orientat spre utilizator.” (European Economic and Social Committee 2017) Portabilitatea datelor ar putea asigura acest deziderat.

AEPD susține promovarea beneficiarilor responsabili și reducerea birocrației în protecția datelor, prin coduri de conduită, audituri, certificări, și o nouă generație de clauze contractuale și

reguli corporative obligatorii. Responsabilitatea beneficiarilor Big Data presupune instituirea unor politici interne și a unor sisteme de control conforme cu legislația în vigoare, prin soluții inteligente și dinamice care să garanteze respectarea principiilor fundamentale (minimizarea datelor, limitarea scopului, calitatea datelor, procesarea corectă și transparentă a datelor, design, limitare de stocare, integritate și confidențialitate).

Etica datelor se bazează pe următoarele principii: **proprietatea** (persoanele fizice dețin propriile date**, transparența tranzacțiilor** (utilizatorii trebuie să aibă acces transparent la proiectarea algoritmului**), consimțământ** (utilizatorul trebuie să fie informat și să își exprime explicit consimțământul cu privire la utilizarea datelor personale**, confidențialitate** (trebuie protejată confidențialitatea utilizatorilor)**, financiar** (utilizatorul să cunoască tranzacțiile financiare rezultate din utilizarea datelor lui personale**), și deschidere** (seturile de date agregate să fie disponibile în mod liber).

Etica în cercetare

Termenul de studiu critic de date (SCD) implică faptul că cercetătorii investighează Big Data din perspective critice. Studierea datelor în acest context implică, pe lângă analiza lor, și încorporarea datelor în practici (cunoașterea), instituții și sisteme politice și economice, prin interacțiunea complexă dintre date și entitățile care le produc, dețin și folosesc.

Un raport al OECD (2013) subliniază că, spre deosebire de normele etice aplicate datelor obișnuite de cercetare, în cazul Big Data: (OECD 2013)

- Colectarea de date nu a făcut obiectul unui proces formal de examinare etică.
- Normele etice obișnuite nu vor fi implementate în cazul Big Data
- Utilizarea datelor pentru cercetare poate să difere de scopul inițial.
- Datele nu mai sunt deținute ca seturi discrete.

Relația dintre cei care oferă datele și cei care le folosesc este adesea indirectă și variabilă. Un raport mai recent al OECD (2016) susține că această relație este mai slabă sau inexistentă, Big Data limitând capabilitățile obișnuite. (OECD 2016)

Stocarea datelor e importantă pentru integritatea cercetării. Datele trebuie să aibă o "proveniență" clară, cu surse și procesare cunoscute, identificate și documentate.

Multe date care nu sunt colectate special pentru cercetare au standarde diferite în cercetarea datelor.

Pentru anumite date, adesea cu valoare comercială (de ex., datele colectate pe Twitter), există restricții legale privind reproducerea lor. (UK Data Service 2017)

Depozitele de date trebuie să respecte standardele de transparență și reproductibilitate.

Conștientizarea

Conștientizarea tipului de date care sunt furnizate în timpul unei înregistrări online (pentru crearea unui cont, sau un abonament, de ex.) este un fapt rar, mai ales că există posibilitatea folosirii unei identități digitale deja existente (profil Facebook, de ex.) în locul unei înregistrări separate, pentru un acces mai rapid. Astfel de situații creează o opacitate cu privire la datele partajate între furnizorul de identitate și serviciul utilizat.

Consimțământul

Pentru utilizarea datelor cu caracter personal ale unei persoane, este nevoie de consimțământul informat și explicit exprimat al acesteia referitor la cine, când, cum și în ce scop se folosesc. Când trebuie partajate datele, aceste utilizări trebuie aduse la cunoștința persoanei. Ar trebui să fie întotdeauna posibilă retragerea consimțământului pentru viitoarele utilizări.

În analizele Big Data, se poate cunoaște foarte puțin despre utilizările viitoare intenționate ale datelor, și despre beneficiile, și riscurile implicate. Aici, există proceduri pentru consimțământul "larg" și "generic" de a împărtăși datele genomice, de ex., și în scopuri diferite. Chiar și atunci când se procedează corect, există anumite provocări practice specifice: obținerea consimțământului în cunoștință de cauză poate fi imposibil sau foarte costisitor, iar valabilitatea consimțământului este disputabil când acordul este obligatoriu pentru a accesa un serviciu.

Controlul

În lumea actuală, datele personale pot fi tranzacționate la fel ca orice monedă în implementarea Big Data. Există opinii diferite în ce măsură această situație este una etică, inclusiv cine să participe la profitul obținut din aceste tranzacționări..

În modelul de tranzacționare a datelor cu caracter personal, transmiterea datelor personale este un cadru care oferă persoanelor posibilitatea de a-și controla identitatea digitală și a crea acorduri granulare de partajare a datelor.

În prezent prinde contur ideea datelor deschise, centrată în jurul argumentului că datele ar trebui să fie disponibile în mod liber. Dorința de a partaja date variază în funcție de persoană.

În cazul copiilor, părinții sau tutorii au responsabilitatea pentru datele lor, care nu pot fi tranzacționate contra beneficii financiare.

La nivel național, un guvern este suveran asupra datelor generate și colectate. La 26 octombrie 2001 a intrat în vigoare Actul Patriot în SUA, iar la 25 mai 2018, Regulamentul general privind protecția datelor 2016/679 (GDPR) la nivelul Uniunii Europene, pentru problemele legate de protecția datelor personale.

În Big Data, relația om-date este asimetrică, bazată pe controlul datelor. ”Dreptul de a fi uitat”, adoptat la nivelul UE, este unul din elementele de bază ale controlului unui individ asupra datelor sale personale.

Transparența

Algoritmii utilizați în Big Data pot determina prejudecăți care afectează sistematic drepturile individului. De aceea, proiectarea algoritmului ar trebui să fie transparentă și inclusivă.

Guvernarea anticipativă implică analize predictive pe baza Big Data pentru a evalua potențiale comportamente, cu implicații etice care pot încuraja prejudecățile și discriminarea.

O persoană care acceptă includerea datelor sale personale în Big Data are dreptul să știe de ce se colectează datele, cum vor fi folosite, cât timp vor fi stocate, și cum pot fi modificate.

Încrederea

Încrederea în sistemele Big Data este legată de interdependențe cu confidențialitatea și conștientizarea. Până în prezent, încrederea a fost considerată din perspectivă strict tehnologică. Se speră să se realizeze arhitecturi hardware și software care ar putea crește încredere între ființe umane și obiecte, și deci o mai mare acceptanță a utilizării datelor personale.

Proprietatea

O întrebare fundamentală în etica cercetării cu Big Data este, cine deține datele? Aceasta implică subiectul drepturilor și obligațiilor asupra proprietății. În legislația europeană, GDPR indică faptul că persoanele dețin propriile date cu caracter personal.

Suma datelor personale ale unui individ formează o identitate digitală.

Protecția drepturilor morale (dreptul de a fi identificat ca sursă a datelor, și de a le controla) ale unui individ se bazează pe opinia că datele personale sunt o expresie directă a personalității acestuia, și nu pot fi transferate unei alte persoane decât, eventual, prin succesiune atunci când individul moare.

Proprietatea implică exclusivitate, respectiv restricționarea implicită a altora în ceea ce privește accesul la proprietate. O proprietate eficientă a datelor personale implică portabilitatea, posibilitatea de a folosi alternative fără a pierde din date. Standardizarea ar ajuta, de asemenea, la curățare datelor personale.

În mod efectiv, în prezent, datele sunt deținute de proprietarul senzorilor, cel care efectuează înregistrarea sau entitatea care deține senzorul.

În UE, s-a restrâns în mod progresiv posibilitatea ca datele cetățenilor UE să fie stocate în afara așa-numitului "Euro cloud", dar nu s-a rezolvat problema datelor deja stocate și prelucrate în altă parte, și ”nu rezolvă dilema etică a modului în care proprietatea asupra datelor este definită în mod filosofic, înainte de a trece la o abordare mai degrabă a legii și a elaborării politicilor.” (European Economic and Social Committee 2017)

Supravegherea și securitatea

Din ce în ce mai multe surse de date sunt disponibile cu ajutorul tehnologiilor avansate, precum circuitele CCTV, GPS, dispozitive mobile, carduri de credit, ATM. De asemenea, supravegherea activă este o metodă de colectare a datelor, dar în același timp de limitare a libertăților cetățenilor. O astfel de supraveghere permanentă determină creșterea stresului oamenilor, și creează tendința acestora de a se comporta într-un anumit mod care să se conformeze normelor așteptate.

Identitatea digitală

Identitatea digitală are avantajul accesului rapid la conținutul online și serviciile conexe. Utilizarea identității digitale are potențialul de a genera discriminare bazată pe reprezentarea unei persoane conform datelor ei online, care de multe ori poate să nu corespundă cu situația reală, într-un proces numit ”dictatura datelor” în care "nu mai suntem judecați pe baza acțiunilor noastre, ci pe baza a ceea ce indică toate datele despre noi ca fiind acțiunile noastre probabile", (Norwegian Data Protection Authority 2013) interacțiunea personală nefiind plasată într-un plan secundar.

Realitatea ajustată

Orice interacțiune a noastră cu Internetul implică posibilitatea stocării datelor noastre personale. Prelucrarea și analiza acestor date determină rezultatele personalizate care ne apar ulterior pe Internet, prin rezultate ale căutărilor noastre, afișarea produselor în magazinele online, afișarea reclamelor, etc. Se generează astfel o versiune mai îngustă și mai personalizată a experienței online anterioare a unui utilizator (așa-numitul "balon de filtrare" (Pariser 2011)). Un avantaj este că utilizatorul va găsi rapid ceea ce caută de obicei, dar excluderea anumitor aspecte, perspective și idei poate duce la o restrângere a creativității și dezvoltarea unei atitudini tolerante prin izolarea politică și socială de celelalte aspecte, prin lipsa unor viziuni pluraliste. (Crawford, Gray, and Miltner 2014)

De-anonimizarea

De-identificarea implică ștergerea sau ascunderea elementelor care ar putea identifica imediat o persoană sau organizație. Legislația din diferite țări privind protecția datelor definește tratamente diferite pentru datele identificabile. Identificabilitatea este văzută din ce în ce mai mult ca un continuum, nu un aspect binar. Riscurile de divulgare cresc simultan cu numărul de variabile, de surse de date și cu puterea analizei datelor. Riscurile de dezvăluire pot fi atenuate, dar nu eliminate. De-identificarea rămâne un instrument vital pentru asigurarea utilizării în siguranță a datelor. (UK Data Service 2017)

Informații perfect anonime luate separat, pot fi combinate cu alte date pentru a identifica în mod unic o persoană cu grade diferite de certitudine. Profilarea poate deveni un instrument puternic, ridicând îngrijorări cu privire la gradul în care este permisă intruziunea în viața unui individ, posibilitatea asigurării securității, și supravegherea.

Inegalitatea digitală

Avantajele dimensiunii mari a datelor sunt clare, dar există și opinii conform cărora acumularea de date la o scară uriașă prezintă riscuri specifice. Din această cauză, sunt puține entități care au acces, prin infrastructură și abilități, la sistemele Big Data. În acest context, costurile și abilitățile necesare accesului duce la anumite inegalități digitale specifice abordate de etică.

Confidențialitatea

În tranzacțiile de date este foarte important să se asigure confidențialitatea:

"Nimeni nu va fi supus la interferențe arbitrare cu intimitatea, familia, casa sau corespondența sa și nici la atacuri asupra onoarei și reputației sale. Toată lumea are dreptul la protecția legii împotriva unor asemenea ingerințe sau atacuri. "- Declarația Organizației Națiunilor Unite privind Drepturile Omului, Articolul 12.

În multe țări, monitorizarea publică datelor de către guvern pentru a observa cetățenii necesită o autorizare explicită printr-un proces judiciar adecvat. Confidențialitatea nu este despre păstrarea secretelor, ci despre alegere, drepturile omului, și libertate.

Adesea confidențialitatea este văzută în mod greșit ca o alegere binară între izolare și progres științific. Protejarea identității în date este posibilă tehnologic, de exemplu utilizând criptarea homomorfă și designul algoritmic.

Confidențialitatea ca o limitare a utilizării datelor poate fi, de asemenea, considerată ne-etică, (Kostkova et al. 2016) în special în asistența medicală, dar trebui ținut cont de faptul că este posibilă extragerea valorii datelor fără a compromite intimitatea.

Confidențialitatea este recunoscută ca un drept uman prin numeroase reglementări naționale și internaționale. Confidențialitatea în cercetare se realizează printr-o combinație de abordări: limitarea datelor colectate, anonimizarea acestora; și reglementarea accesului la date. În cazul cercetării Big Data apar probleme specifice: ambiguitatea între termenii "privațiune" și "confidențialitate; declararea spațiilor sociale ca publice sau private; necunoașterea riscurilor de confidențialitate de către utilizatori; distincția neclară între uzanțele publice și private. În prezent există dispute dacă știința datelor ar trebui să fie clasificată ca o cercetare a subiecților umani, și deci nesupusă normelor obișnuite de confidențialitate.

Cercetarea Big Data

Prin noile concepte de "daune algoritmice", "analize predictive", etc., algoritmii folosiți în prezent în operațiunile Big Data depășesc viziunea tradițională a confidențialității. Conform Consiliului Național pentru Știință și Tehnologie,

”"Algoritmii analitici" sunt algoritmi pentru prioritizare, clasificare, filtrare și predicție. Utilizarea acestora poate crea probleme de confidențialitate atunci când informațiile utilizate de algoritmi sunt inadecvate sau inexacte, atunci când apar decizii incorecte, atunci când nu există mijloace rezonabile de recurs, atunci când autonomia unui individ este direct legată

de rezultatul algoritmic sau atunci când folosirea algoritmilor predictivi încurajează alte daune asupra vieții private.” (NSTC (National Science and Technology Council) 2016, 18)

Cercetările Big Data sunt ceea ce eticianul James Moor ar numi "harababura conceptuală" datorită ”incapacității de a conceptualiza în mod corect valorile etice și dilemele de joc într-un context tehnologic nou.” (Buchanan and Zimmer 2018) În această situație confidențialitatea este asigurată printr-o combinație de diferite tactici și practici (medii controlate sau anonime, limitarea informațiilor personale, anonimizarea datelor, restricții de acces, securizarea datelor, etc.). În general, toate noțiunile conexe devin confuze în cazul Big Data. Astfel, postările sociale sunt considerate publice în rețelele sociale în cazul unei setări corespunzătoare. Dar rețelele sociale sunt medii complexe de interacțiuni socio-tehnice unde utilizatorii nu înțeleg întotdeauna funcționalitatea setărilor și termenii de utilizare. Astfel, există o incertitudine în ceea ce privește intențiile și așteptările utilizatorilor, iar aceste deficiențe conceptuale în contextul cercetărilor Big Data conduc la incertitudini în ceea ce privește necesitatea consimțământului informat.

Bibliografie

Buchanan, Elizabeth A., and Michael Zimmer. 2018. “Internet Research Ethics.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2018. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2018/entries/ethics-internet-research/.

Crawford, Kate, Mary L. Gray, and Kate Miltner. 2014. “Big Data| Critiquing Big Data: Politics, Ethics, Epistemology | Special Section Introduction.” *International Journal of Communication* 8 (0): 10. https://ijoc.org/index.php/ijoc/article/view/2167.

DG Connect. 2015. “Study on Personal Data Stores Conducted at the Cambridge University Judge Business School.” Text. Digital Single Market - European Commission. August 7, 2015. https://ec.europa.eu/digital-single-market/en/news/study-personal-data-stores-conducted-cambridge-university-judge-business-school.

European Economic and Social Committee. 2017. “The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context.” European Economic and Social Committee. February 22, 2017. https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/ethics-big-data.

Kostkova, Patty, Helen Brewer, Simon de Lusignan, Edward Fottrell, Ben Goldacre, Graham Hart, Phil Koczan, et al. 2016. “Who Owns the Data? Open Data for Healthcare.” *Frontiers in Public Health* 4. https://doi.org/10.3389/fpubh.2016.00007.

MORO, Veronica. 2016. “Meeting the Challenges of Big Data.” Text. European Data Protection Supervisor - European Data Protection Supervisor. November 16, 2016. https://edps.europa.eu/data-protection/our-work/publications/opinions/meeting-challenges-big-data_en.

Norwegian Data Protection Authority. 2013. “Big Data – Privacy Principles under Pressure.” https://www.datatilsynet.no/globalassets/global/english/big-data-engelsk-web.pdf.

NSTC (National Science and Technology Council). 2016. "National Privacy Research Strategy." https://obamawhitehouse.archives.gov/sites/default/files/nprs_nstc_review_final.pdf.

OECD. 2013. "New Data for Understanding the Human Condition: International Perspectives." http://www.oecd.org/sti/inno/new-data-for-understanding-the-human-condition.pdf.

———. 2016. "Research Ethics and New Forms of Data for Social and Economic Research," November. https://doi.org/10.1787/5jln7vnpxs32-en.

Pariser, Eli. 2011. *The Filter Bubble: What The Internet Is Hiding From You*. Penguin Books Limited.

UK Data Service. 2017. "Big Data and Data Sharing: Ethical Issues." https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf.